

# Characterizing intimate mixtures of materials in hyperspectral imagery with albedo-based and kernel-based approaches

Robert S. Rand

National Geospatial-Intelligence Agency (NGA), Springfield, VA 22150

Ronald G. Resmini

The MITRE Corporation, McLean, VA 22102

David W. Allen

The National Institutes of Standards and Technology

Gaithersburg, MD 20899

## ABSTRACT

Linear mixtures of materials in a scene often occur because the pixel size of a sensor is relatively large and consequently they contain patches of different materials within them. This type of mixing can be thought of as areal mixing and modeled by a linear mixture model with certain constraints on the abundances. The solution to these models has received a lot of attention. However, there are more complex situations, such as scattering that occurs in mixtures of vegetation and soil, or intimate mixing of granular materials like soils. Such multiple scattering and microscopic mixtures within pixels have varying degrees of non-linearity. In such cases, a linear model is not sufficient. Furthermore, often enough, scenes may contain cases of both linear and non-linear mixing on a pixel-by-pixel basis. This study considers two approaches for use as generalized methods for un-mixing pixels in a scene that may be linear (areal mixed) or non-linear (intimately mixed). The first method is based on earlier studies that indicate non-linear mixtures in reflectance space are approximately linear in albedo space. The method converts reflectance to single-scattering albedo (SSA) according to Hapke theory assuming bidirectional scattering at nadir look angles and uses a constrained linear model on the computed albedo values. The second method is motivated by the same idea, but uses a kernel that seeks to capture the linear behavior of albedo in non-linear mixtures of materials. The behavior of the kernel method can be highly dependent on the value of a parameter, gamma. Furthermore, both methods are dependent on the choice of endmembers, and also on RMSE (root mean square error) as a performance metric. This study compares the two approaches and pays particular attention to these dependencies. Both laboratory and aerial collections of hyperspectral imagery are used to validate the methods.

**Keywords:** Hyperspectral, kernel methods, single scattering albedo, Hapke theory, linear spectral mixing, non-linear spectral mixing.

## 1. INTRODUCTION

In a broad sense, spectral mixing phenomenology in multispectral and hyperspectral imagery can be treated in two ways, depending on how the materials in a scene are presumed to be mixed. A linear mixing model is usually suitable in cases where materials are presumed to be non-overlapping (areal) and can be mathematically expressed as a linear combination, where the weights in the combination are associated with the abundances of each material. The endmembers are spectra (hopefully) representing unique materials in a given image such as water, soil, and vegetation. Abundances are the percentage of each endmember within a given pixel. On the other hand, it is well known that intimately mixed materials frequently exhibit nonlinear spectral mixing. Granular materials, such as soils, are often intimate mixtures of numerous different inorganic (minerals) and organic (humic) substances. And since soils are often significant constituents of spectral remote sensing scenes, intimate mixing may safely be assumed to be a common phenomenon. Another case is the mixing of fluids, such as oil and water, as occurred during the Deepwater Horizon oil spill disaster a few years ago. We will consider this second case in more detail, below.

Linear mixing is modeled as linear combination of the spectra from multiple endmembers. The problem is posed in a number of ways using theory from Linear Statistical Models with variations that impose either physical or sparseness constraints.<sup>1-7</sup> Sites in a scene are labeled as containing one, two, or perhaps many endmembers. Some approaches begin

with what is sometimes referred to as a full model, where the full model contains all possible variables (endmembers), and subsequently eliminates variables that do not contribute to the statistical significance (e.g., using an F-statistic) of the model.<sup>8</sup> Alternatively, other approaches are synonymous with step-wise regression, where the process begins with a pair of variables (endmembers) and adds variables if they contribute significantly to the model.<sup>6,8</sup>

Intimate mixing is modeled as non-linear combinations of spectra from multiple endmembers. An intimate mixture model can be described by nonlinear functions, which are justified by Hapke scattering theory<sup>9</sup> and photometric phase functions.<sup>18</sup> This approach can be used to convert reflectance to single scattering albedo (SSA). An exact model can be difficult to obtain, but can show up to 30% improvement in measurements over the linear mixing model when intimate mixtures are present.<sup>10</sup> Many years ago studies showed success in using a Constrained Energy Minimization (CEM) method and other linear methods applied to SSA data.<sup>20-22</sup>

As an alternative to computing SSA, kernel functions have been introduced as a way to generalize linear algorithms to nonlinear data.<sup>11,12</sup> In the case of detection and classification applications, kernel functions can induce high dimensional feature spaces. In these spaces, previously non-separable classes can be made linearly separable. Thus, linear methods can be applied in this new feature space that provides nonlinear boundaries back in the original data space. Another example is the kernel PCA method.<sup>13</sup> The kernel, in this case, is not used to induce a high dimensional space, but is used to better match the data structure through nonlinear mappings. It is in this mode that kernels can be used to produce nonlinear mixing results while essentially using a linear mixture model. What is more appealing is that the physics suggest that such a method is ideal if one can model the kernel correctly.

The drawback with the earlier kernel algorithms for classification and detection is that they produced abundance estimates that do not meet the non-negativity and sum-to-one constraints. This was solved by the development of a Kernel Fully Constrained Least Squares (KFCLS) which computes kernel based abundance estimates to meet the physical abundance constraints.<sup>14</sup> Further investigation of the KFCLS method has resulted in (1) the development of a generalized kernel for areal (linear) and intimate (non-linear) mixtures<sup>15</sup> and (2) an adaptive kernel-based technique for mapping areal and intimate mixtures.<sup>16</sup> The generalized kernel and adaptive technique provides a way to adaptively estimate a mixture model suitable to the degree of non-linearity that may be occurring at each pixel in a scene. This is important because a scene may contain both areal and intimate mixtures and we don't always know a priori which model is appropriate on a pixel-by-pixel basis. These methods were developed further by Broadwater and Banerjee.<sup>17,19</sup> Based on their work, a study investigating the behavior of the generalized KFCLS and adaptive kernel-based techniques was recently performed using both user-defined and SVDD automatically-generated endmembers.<sup>23</sup>

Recently a study using laboratory data was performed comparing the performance of the generalized KFCLS applied to reflectance spectra with the FCLS method applied to spectra converted to SSA.<sup>24</sup> One of the conclusions of this effort was that similar accuracy in abundance estimates can be achieved using the SSA-based method, but with much faster computation time.

Overall, the impact of nonlinear spectral mixing on algorithm results must be well understood if we are to achieve a major goal of hyperspectral imaging (HSI): the areal mapping and quantification of materials that comprise remotely sensed scenes. In this study we attempt to further this understanding through investigating both phenomenology-based Single Scattering Albedo as well as mathematical-based kernel methods. Our current effort continues this investigation on laboratory and airborne imagery data. The laboratory experiment is performed on a highly controlled data containing pre-determined non-linear mixtures of two materials. The airborne experiment is performed on an AVIRIS scene collected over the oil spill region in the Gulf of Mexico during the Deepwater Horizon oil incident.

## 2. DESCRIPTION OF ALGORITHMS

### 2.1 Fully Constrained Least Squares

What is often called the Fully Constrained Least Squares (FCLS)<sup>7</sup> spectral mixing model can be expressed as

$$\mathbf{x} = \mathbf{E}\mathbf{a}, \quad \text{with constraints } a_i \geq 0 \text{ and } \sum_{i=1}^N a_i = 1 \quad (1)$$

where  $\mathbf{x}$  is an  $L \times 1$  vector containing the spectral signature of the current image pixel,  $\mathbf{E}$  is an  $L \times N$  matrix containing the endmember signatures (the  $i^{\text{th}}$  column contains the  $i^{\text{th}}$  endmember spectrum), and  $\mathbf{a}$  is an  $N \times 1$  vector containing the estimated abundances (the  $i^{\text{th}}$  entry represents the abundance value  $a_i$ ).  $N$  is the number of endmembers within the image and  $M$  is the number of endmembers within the scene.

The FCLS method has been quite successful in the past for modeling linear mixing phenomenology. In our present study we will be using FCLS in two ways. The method will be used as a benchmark to compare against with the proposed non-linear methods. The method will also be used in one of the non-linear approaches, where we will apply the FCLS to spectra that has been converted to Single Scattering Albedo (SSA), as discussed immediately, below.

### 2.2 Using Hapke Theory: Fully Constrained Least Squares (FCLS) applied to Single Scattering Albedo Spectra.

As just mentioned, previous studies indicate that intimate (non-linear) mixtures in reflectance space, are approximately linear in albedo space. Therefore, rather than apply the complex computations involved with (5) on reflectance data to estimate abundances of intimate mixtures, why not apply a linear mixing method on albedo? We consider this by applying the FCLS method on data that's been converted to SSA. Conversion to SSA is described in Resmini et al.<sup>20</sup> (1996) and Resmini<sup>21</sup> (1997) (both studies following Hapke<sup>9</sup> (1993); and Mustard and Pieters<sup>18</sup> (1987)) assuming the reflectance spectra are bidirectional. SSA spectra were also generated assuming the input reflectance spectra are hemispherical-directional. The expressions to transform reflectance spectra to SSA are given by eqs. (2) and (3) for bi-directional (bd) reflectance and for hemispherical-directional (hd) reflectance, respectively. In the derivation of both expressions, phase angle is large enough that the opposition effect is assumed negligible.

$$\bar{\omega} = 1.0 - \left( \frac{[(\mu_0 + \mu)^2 \Gamma^2 + (1.0 + 4.0\mu\mu_0\Gamma)(1.0 - \Gamma)]^{0.5} - (\mu_0 + \mu)\Gamma}{(1.0 + 4.0\mu\mu_0\Gamma)} \right)^2 \quad (2)$$

$$\bar{\omega} = 1.0 - \left( \frac{1.0 - \Gamma}{(1.0 + 2.0\mu\Gamma)} \right)^2 \quad (3)$$

In (2) and (3),  $\bar{\omega}$  is single scattering albedo;  $\Gamma$  is reflectance factor (Hapke<sup>22</sup>),  $\mu_0$  is the cosine of the angle of incidence of the illumination, and  $\mu$  is the cosine of the viewing angle. Note that one reflectance is calculated as described previously in Section 2.1 though two different equations are used to generate the two sets of SSA spectra.

In the experiments that follow, we refer to this approach as the "FCLS on SSA" method, or as simply the ALBEDO method. The conversion of reflectance spectra to SSA using (2) and (3) is very fast as compared to the GKLS method.

### 2.3 Generalized Kernel Fully Constrained Least Squares

A kernel-based mixing model was derived previously by Broadwater, Chellappa, and Banerjee,<sup>14</sup> where the method estimates the abundances of a mixture using the expression

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \frac{1}{2} \left( K(\mathbf{x}, \mathbf{x}) - 2\hat{\mathbf{a}}^T K(\mathbf{E}, \mathbf{x}) + \hat{\mathbf{a}}^T K(\mathbf{E}, \mathbf{E}) \hat{\mathbf{a}} \right), \quad \text{s.t. } a_i \geq 0, \quad \forall_i \quad (4)$$

where  $\mathbf{E}$  is the endmember matrix and  $\hat{\mathbf{a}}$  is the estimated abundance vector, discussed above for (1). The abundance estimates are computed using a quadratic programming method to enforce the non-negativity constraint. The specific choice of kernel determines the ability of this method to respond to different types of mixing. Choosing the linear kernel  $K(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$  is ideal for modeling linear mixture; however, it is not a suitable kernel for intimate mixtures. A physics-inspired kernel has also been proposed and was shown to provide significantly improved behavior to model non-linear mixtures.<sup>17</sup> However, a conclusion of that effort was that although each kernel provides good results for the type of mixing intended, only one kernel or the other could be used, for either areal mixtures or intimate mixtures, but not both.

The approach was further developed by Broadwater and Banerjee into a generalized method for adaptive areal and intimate mixtures.<sup>15,16</sup> This method is motivated by attempting to simulate Hapke theory for Single Scattering Albedo (SSA), making use of the kernel:

$$K_\gamma(\mathbf{x}, \mathbf{y}) = \left(1 - e^{-\gamma \mathbf{x}}\right) \left(1 - e^{-\gamma \mathbf{y}}\right) \quad (5)$$

The kernel in (5) can model either areal or intimate mixtures so long as the correct  $\gamma$  is used. If  $\gamma$  is very small then  $K_\gamma(\mathbf{x}, \mathbf{y})$  approximates linear mixing. If  $\gamma$  is large then  $K_\gamma(\mathbf{x}, \mathbf{y})$  approximates intimate mixing in cases when the reflectance occurring from intimate mixing is modeled as

$$\mathbf{r} = \frac{\mathbf{w}}{4(\mu + \mu_0)} \left[ H(\mathbf{w}, \mu) H(\mathbf{w}, \mu_0) \right] \quad (6)$$

where  $\mathbf{r}$  is the reflectance vector,  $H$  is Chandrasekhar's function for isotropic scattering,  $\mathbf{w}$  is the average single-scattering albedo vector,  $\mu_0$  is the cosine of the angle of incidence, and  $\mu$  is the cosine of the angle of emergence.<sup>18</sup> The computation is similar in form to (4) except the minimization is done according to

$$\hat{\gamma} = \arg \min_{\gamma} \frac{1}{2} \left( K_\gamma(\mathbf{x}, \mathbf{x}) - 2\hat{\mathbf{a}}_\gamma^T K_\gamma(\mathbf{E}, \mathbf{x}) + \hat{\mathbf{a}}_\gamma^T K_\gamma(\mathbf{E}, \mathbf{E}) \hat{\mathbf{a}}_\gamma \right), \quad \text{s.t. } a_i \geq 0, \quad \forall_i \quad (7)$$

where  $\hat{\mathbf{a}}_\gamma$  is the abundance estimate and  $K_\gamma(\mathbf{x}, \mathbf{y})$  is the kernel evaluated with the parameter value  $\gamma$ . A numerical optimization based on the golden search method is used to minimize (7).<sup>19</sup> An implementation of this generalized method is investigated, which we refer to as Generalized Kernel Least Squares (GKLS).

In addition to a fixed-gamma GKLS implementation, an automated GKLS method is investigated that attempts to select the most appropriate gamma based on achieving a minimum of the model's RMSE. In this manner, at least theoretically, the automated GKLS method has the ability to respond differently to differing degrees of non-linearity to select the best gamma and compute more precise estimates of abundance.

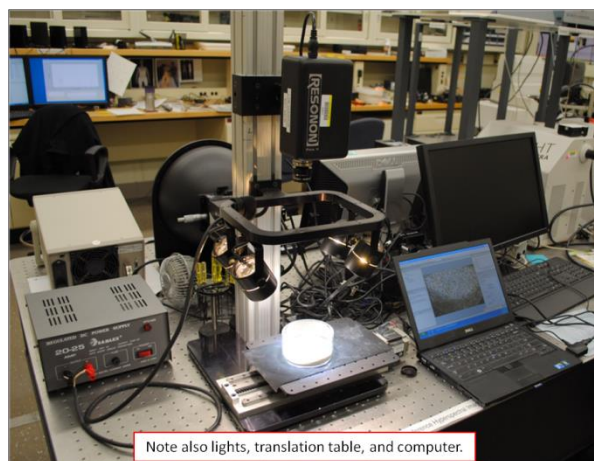
### 3. DESCRIPTION OF EXPERIMENTS

#### 3.1. Overview of Experiments

Two experiments are performed in this study. The first is an experiment conducted in the laboratory. In this experiment, two granular materials were custom fabricated and mechanically mixed to form intimate mixtures. The materials are spherical beads of didymium glass and soda-lime glass, both ranging in particle size from 63  $\mu\text{m}$  to 125  $\mu\text{m}$ . The mixtures, which exhibit largely nonlinear spectral mixing, were then observed with a visible/near-infrared (VNIR; 400 nm to 900 nm) hyperspectral imaging (HSI) microscope. The second experiment is performed using airborne hyperspectral imagery, an AVIRIS scene that was collected over the oil spill region in the Gulf of Mexico during the Deepwater Horizon oil incident. The SSA, FCLS, and GKLS methods are applied to the data experiments.

#### 3.2. Experiment 1 - Laboratory Experiment

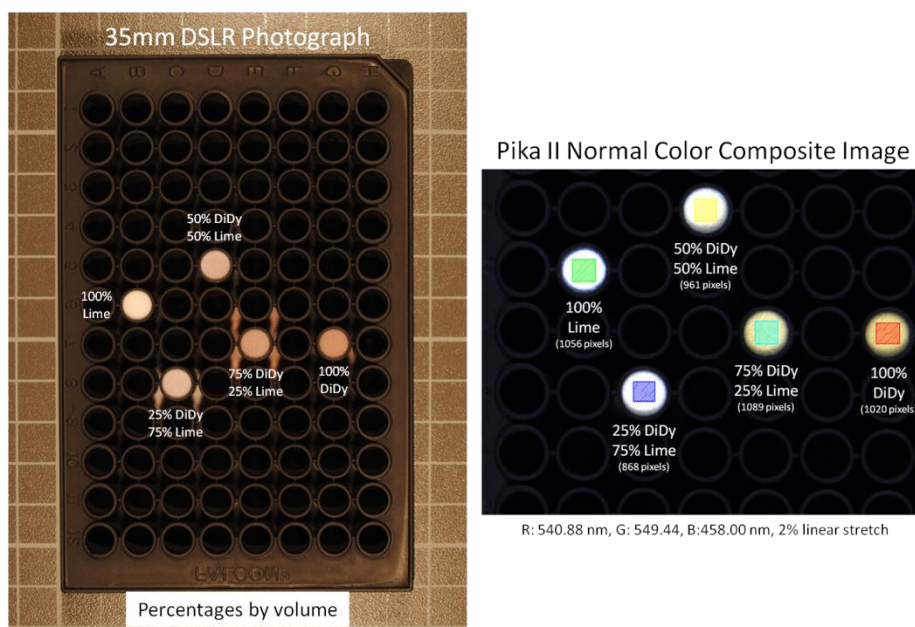
A Resonon Pika II imaging spectrometer with a Xenoplan 1.4/23-0902 objective lens shown in Figure 1 is used to measure the glass bead mixtures.<sup>29,20,31</sup> The Pika II is mounted nadir-looking at a mechanical translation table on which the sample to be imaged is placed. The Pika II is a pushbroom sensor with a slit aperture, thus the need for a translation table to move the sample to facilitate hyperspectral image cube formation. The height of the sensor above the table is user selectable; a height was chosen such that all mixtures are captured in the same scene thus the data have a ground sample distance of  $\sim 75 \mu\text{m}/\text{pixel}$ . Though capable of acquiring 240 bands from 400.0 nm to 900.0 nm, the sensor was configured to acquire 80 bands by binning spectrally by three resulting in a sampling interval of  $\sim 6.25 \text{ nm}$  and high signal-to-noise ratio spectra. Four quartz-tungsten-halogen (QTH) lamps are used for illumination approximating a hemispherical-directional illumination/viewing geometry. Sensor and translation table operation, data acquisition, and data calibration are achieved by software that runs on a laptop computer. Calibration consists of a measurement of dark frame data (i.e., acquiring a cube with the lens cap on) and a measurement of a polytetrafluoroethylene (PTFE) reference plaque (large enough to entirely fill the field-of-view). Then, for each HSI cube measured, the sensor's software first subtracts the dark data and then uses the PTFE data (also dark subtracted) to ratio the spectral measurements to give relative reflectance (also known as reflectance factor; Hapke, 1993<sup>29</sup>; Schott, 2007<sup>32</sup>).



**Figure 1:** Photograph of the VNIR HSI microscope. The Resonon Pika II is shown with the Xenoplan 1.4/23-0902 objective lens [33].

Three binary mixtures (and the two endmembers) are constructed and emplaced in the wells of a 96-well sample plate: 0%/100%, 25%/75%, 50%/50%, 80%/20%, and 100%/0% of didymium/soda-lime (percentages by volume). This was done as follows: Five cells of a 96-well sample plate, spray-painted flat black, were filled with the various glass bead mixtures; this is shown in Figure 2. The volume of each cell is 330  $\mu\text{L}$  (0.33 mL). Three binary mixtures and the two 100% endmembers are constructed and emplaced in five of the wells of the 96-well sample plate: 0%/100%, 25%/75%, 50%/50%, 80%/20%, and 100%/0% didymium/soda-lime. The materials are spherical beads of didymium glass and soda-lime glass both ranging in particle size from 63  $\mu\text{m}$  to 125  $\mu\text{m}$ . Mixtures are by volume. This is a data set with only non-linear spectral mixing; the glass beads, didymium and soda-lime, are translucent. Their chemical composition, densities,

and particle size range are well known. Note that the glass bead particle size range is much larger than the VNIR wavelengths used in this analysis. The glass beads and their mixtures display subtle, though interesting, gonioapparent changes in color.

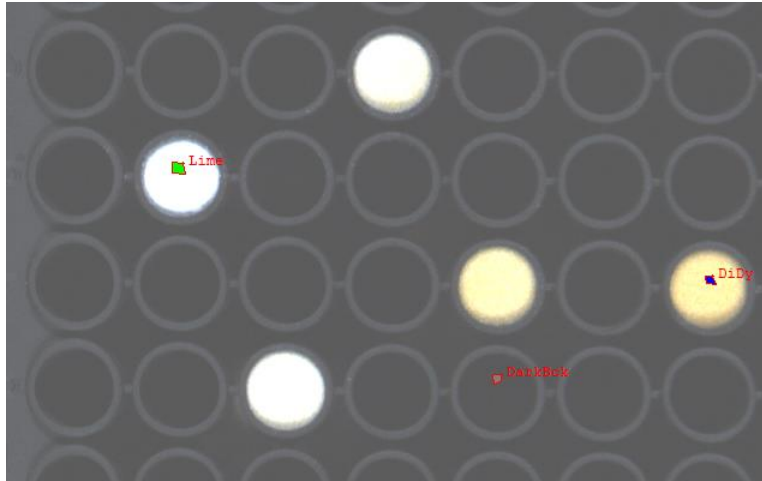


**Figure 2:** (left) 35 mm digital single-lens reflex (DSLR) camera photograph of the 96-well plate containing the glass beads. (right) A Pika II normal color composite image (2% linear stretch of the bands used in the red-green-blue [RGB] image). All percentages are by volume of glass bead type indicated. Spillage onto the plate is evident in the photo on the left but not in the image on the right. Changing the stretch of the Pika II imagery will reveal the spillage [33].

Fifteen data cubes were acquired; however, we focus here on the analysis of one cube comprised of 640 samples, 500 lines, and 75 bands ranging from 434.0 nm to 885.0 nm. Of the 80 bands acquired the first 5 were discarded due to noise content.

Training and test data were extracted from the selected hyperspectral cube. Figure 3 shows polygons defining the training and test regions drawn on top of a Red-Green-Blue (RGB) color composite of this cube. For training, three training endmembers are defined: DiDy (100%), Lime (100%) and Background. The spectra within the small training polygon regions were extracted and the averages of the spectra the regions were used as endmember spectra for the three methods under investigation. For purposes of testing the performance of the algorithm, five regions were extracted, corresponding to the five mixtures 100% DiDy, 75/25% DiDy/Lime, 50/50% DiDy/Lime, 25/75% DiDy/Lime, and 0/100% DiDy/Lime. Two additional test regions of background spectra were extracted. The training regions are shown in Figure 3a and the test regions are shown in Figure 3b. Note that none of these test regions overlapped the training regions. The spectra of the training endmembers are shown in Figure 4.

The image-derived training endmembers, as just described, are used to investigate the three methods described in Section 2.0: (1) FCLS applied in reflectance space; (2) GKLS applied in reflectance space; and (3) “FCLS applied in SSA space. For purposes of conciseness in reporting results, these methods henceforth will be referred to as the FCLS, GKLS, and SSA methods, respectively.



(a) RGB Composite of the HSI cube showing training regions



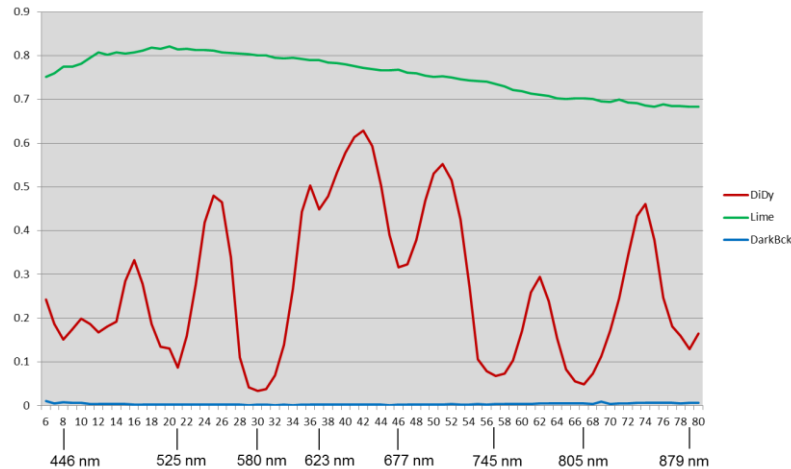
(b) RGB Composite of the HSI cube showing test regions

**Figure 3:** RGB composite images of the hypercube showing the training and test regions used in the experiment trials.

Numerous factors affect the performance of the methods. Three factors affecting the performance of all the methods are (1) the number the endmembers used in a model; (2) how well these endmembers span the space in which the mixing occurs; and (3) the Root Mean Square Error (RMSE) threshold for eliminating bad fits between the observed and model-estimated spectra. In addition, the GKLS method uses a kernel parameter “gamma” that determines the non-linear behavior introduced by the models’ kernel. We test the GKLS method at fixed values of gamma: 0.1, 0.5, 1.0, 2.0, 3.0, 4.0, 5.0 and 6.0, as well as the automated GKLS. For the SSA method, performance might be affected by the type SSA conversion: Hemispherical or Bi-Directional. If Bi-Directional, the input and output angles are other factors. In our case, we report on the results for an SSA conversion made assuming bi-directional reflectance with nadir input and output angles. We also tried the SSA conversions at other angles, but we didn’t notice any noteworthy difference.

The experiment trials were made on the entire scene. Both qualitative results (shown by pictures of the entire scene) and quantitative results (applied in the test regions) are given.





**Figure 4:** Mean spectra for the three endmembers (DiDy, Lime, and DarkBackground) are shown.

### 3.3. Experiment 2 - Airborne Experiment

An airborne experiment is performed using a scene extracted from AVIRIS data collected during a JPL/USGS campaign in response to the Deep Water Horizon (DWH) oil spill incident.<sup>27</sup> Out of the numerous flight lines of AVIRIS acquired on May 17, 2010, Run 11 of AVIRIS was selected, which includes the incident site of the oil spill. This run was collected at an altitude of 28,000 feet resulting in a Ground Sampling Distance (GSD) of approximately 9 meters. The full size of the imagery is 677 pixels by 16,835 lines with 224 bands of reflectance data in the range 0.365-2.49 microns. As part of the calibration to reflectance, the imagery was scaled by a factor of 20,000. For Experiment 2, a portion of this scene of size “500 samples by 500 lines” was extracted for processing.

Figure 5 shows a true-color RGB composite of this scene. The scene is mostly water with large portions containing oil. Different states of the oil can be seen visibly, appearing in different colors ranging from almost black, to dark brown, to light brown, and to orange. This small scene contains a great diversity of oil thicknesses.

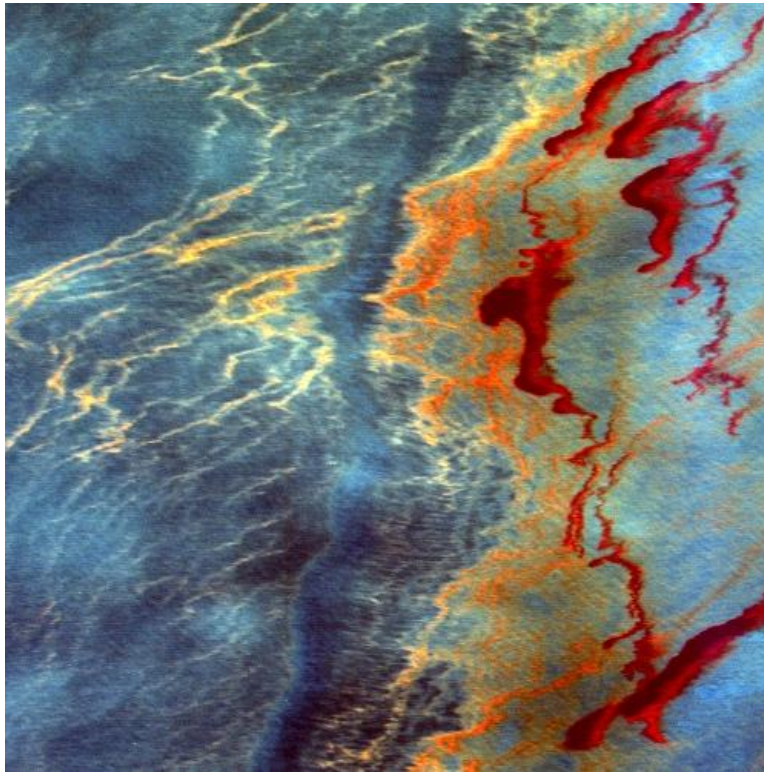
Although precise ground truth is not available due to the dynamically changing environment of an oil spill on water, previous studies using feature-based spectroscopy methods determined the locations of three predominant oil thicknesses in the region. As a part of these studies, during the DWH campaign, USGS collected numerous oil samples and brought them back to their laboratory for measurements.<sup>27</sup>

Figure 6 shows laboratory-measured spectra of three oil samples at differing thicknesses. Absorption features due to C-H with band centers (bc) at 1.2 microns and 1.73 microns are highlighted in red and green, respectively. The left edge of an absorption feature is labeled *lc* (left continuum edge) and the right edge is labeled *rc* (right continuum edge). The blue-colored diagonal line shows the inferred continuum spectra  $C_s$  of the 1 mm thickness oil sample. The edges and center for the feature near 1.2 microns are shown in the figure. Similar edges and center are defined near the 1.7 micron feature, but are not depicted

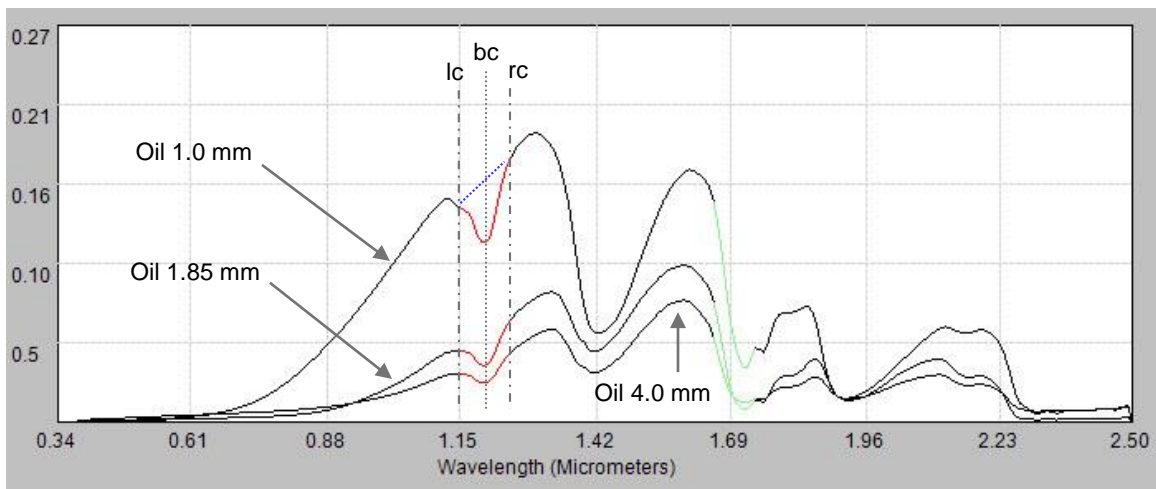
For purposes of training and defining the endmembers, laboratory spectra for pure oil samples with 4mm thickness and 1.85 mm thickness are used, as well as an oil/water emulsion having a thickness of 1.0 mm. Because much of the emulsion behavior is difficult to model in the laboratory, user-defined image spectra are also used: orange colored spectra (in the RGB true-color composite) and water spectra with trace oil.

Trials are performed for the FCLS, GKLS, and SSA methods. Based on the results from the laboratory experiment, we limited our testing of the GKLS method to a fixed value of gamma ( $G = 5$ ). We report on the results for an SSA conversion made assuming bi-directional reflectance with nadir input and output angles.





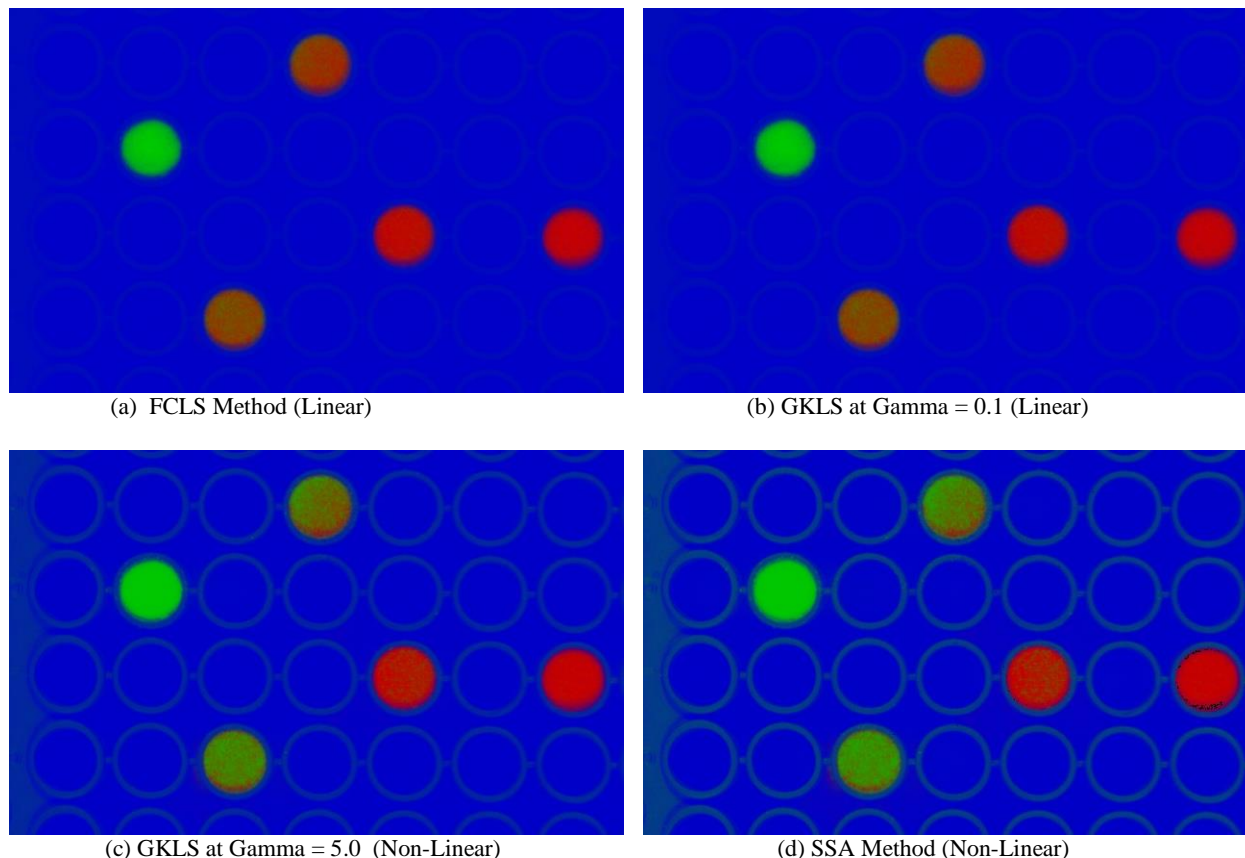
**Figure 5:** An RGB composite scene (500 samples by 500 lines) from a portion of AVIRIS Run 11 is shown in (a). [Source of imagery: U.S. Geological Survey]



**Figure 6:** Spectra of three oil samples at differing thicknesses are shown. Absorption features due to C-H with band centers (bc) at 1.2 microns and 1.73 microns are highlighted in red and green, respectively. The left edge of an absorption feature is labeled *lc* (left continuum edge) and the right edge is labeled *rc* (right continuum edge). The blue-colored diagonal line shows the inferred continuum spectra  $C_s$  of the 1 mm thickness oil sample.

## 4. RESULTS

The results for Experiment 1 (the Laboratory Experiment) are shown in Figures 7 to 9, as well as Tables 1 and 2. Figure 7 shows RGB Color Abundance Maps for four of the trials (Red=DiDy, Green=Lime, Blue= Background). Qualitatively, Figures 7a and 7b show poor correspondence to the known mixtures shown in Figure 2. Figures 7c and 7d show much better correspondence to the known mixtures shown in Figure 2. The variations in color within the discs containing the three mixtures (75/25, 50/50/ and 25/75) is noteworthy. These variations indicate the methods are detecting notable variance in the abundance proportions. This is in spite of the experiment goals to prepare mixture proportions that are as uniform as possible. There is little reason to doubt that these variations are real and that the methods (particularly, GKLS and SSA) are responding correctly. Static clinging and other inter-particle interactions can easily account for the clumping and variations observed.

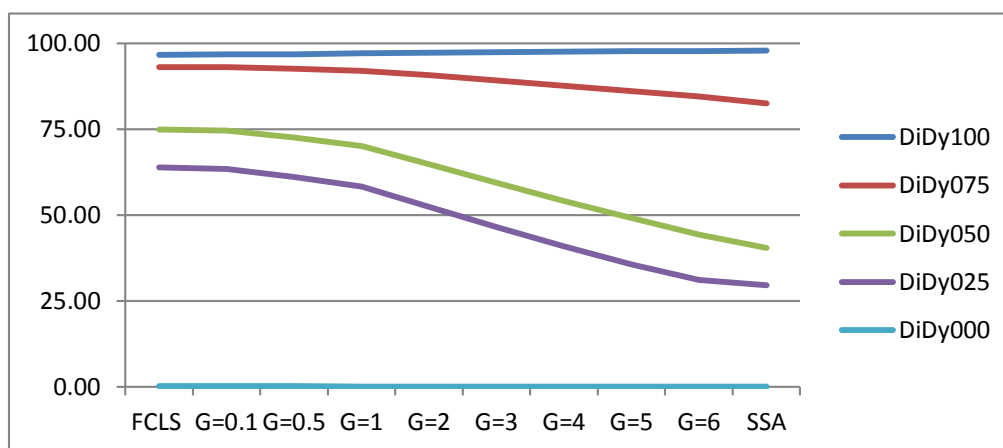


**Figure 7:** RGB composite images showing Color Abundance Maps for four of the trials (Red=DiDy, Green=Lime, Blue= Background). Figures 7a and 7b show poor correspondence to the known mixtures shown in Figure 2. Figures 7c and 7d show much better correspondence to the known mixtures.

Table 1 lists the average estimated abundances for the FCLS, GKLS, and SSA methods in the five test regions. Figure 8 shows these results graphically. The “truth” (actual physically measured) percent by volume of DiDy for these regions varied slightly from the goal of 100%, 75%, 50% 25% and 0%. In reality these were measured as 100%, 78.8%, 50.5%, 24.2%, and 0% for DiDy100, DiD075, DiDy050, DiDy025, and DiDy000, respectively. Results for the GKLS method using a fixed gamma of  $G=0.1$ ,  $G=0.5$ ,  $G=1.0$ ,  $G=2.0$ ,  $G=3.0$ ,  $G=4.0$ ,  $G=5.0$ , and  $G=6.0$  are given. This table shows FCLS to be poor at predicting the abundances for DiDy075 (93.11% vs. 78.8%) and DiDy050 (74.97% vs. 50.5%), as well as very poor at predicting DiDy025 (63.88% vs. 24.2%). The results of GKLS for small gamma agree with theoretical expectations of approximately a linear model. Specifically, the predication of GKLS at  $G=0.1$  is almost exactly the same as the FCLS method. Out of the eight gamma values tested, GKLS at  $G=5.0$  provides the closest prediction for DiDy50 (49.08% vs. 50.5%) and is only slightly worse at predicting the correct abundance than GKLS at  $G=6.0$ .

**Table 1. Abundance Results:** The average estimated abundances are listed for the FCLS, GKLS, and SSA methods in the five test region. The “truth” for these regions (actual physically measured proportions by volume) is 1.0, 0.788, 0.505, 0.242, and 0.0 for DiDy100, DiDy075, DiDy050, DiDy025, and DiDy000, respectively. Results for the GKLS method using a fixed gamma of G= 0.1, G=0.5, G=1.0, G=2.0, G=3.0, G=4.0, G=5.0, and G=6.0 are given, as well as for the automated GKLS.

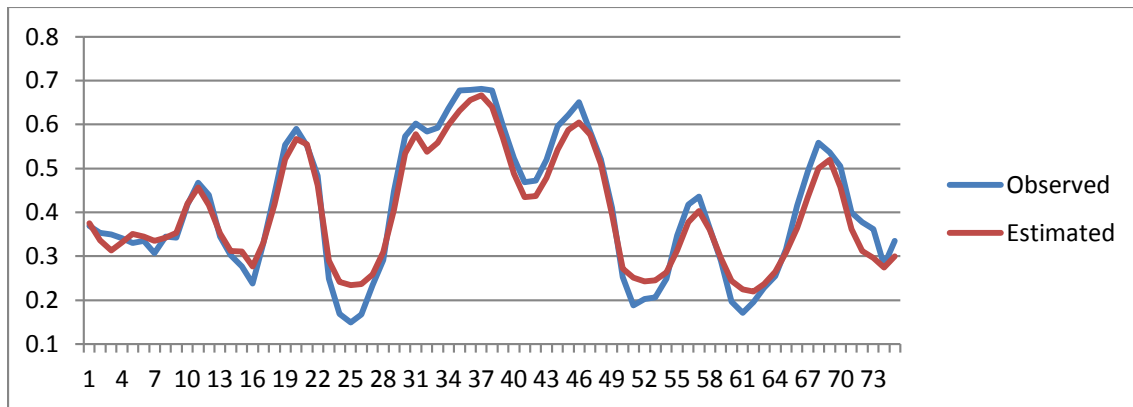
	<u>FCLS</u>	<u>G=0.1</u>	<u>G=0.5</u>	<u>G=1</u>	<u>G=2</u>	<u>G=3</u>	<u>G=4</u>	<u>G=5</u>	<u>G=6</u>	<u>G=Auto</u>	<u>SSA</u>
<b>DiDy100</b>	0.97	0.97	0.97	0.97	0.97	0.97	0.98	0.98	0.98	0.97	0.98
<b>DiDy075</b>	0.93	0.93	0.93	0.92	0.91	0.89	0.88	0.86	0.84	0.90	0.82
<b>DiDy050</b>	0.75	0.75	0.73	0.70	0.65	0.59	0.54	0.49	0.44	0.55	0.40
<b>DiDy025</b>	0.64	0.63	0.61	0.58	0.52	0.47	0.41	0.36	0.31	0.33	0.30
<b>DiDy000</b>	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00



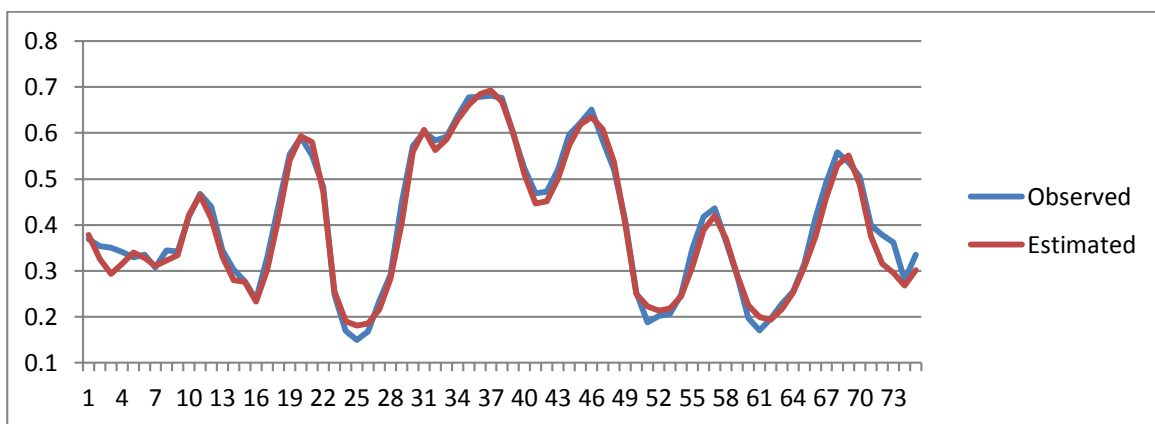
**Figure 8:** A graph of the results listed in Table 1 for the FCLS, GKLS, and SSA methods is displayed.

**Table 2. Model Diagnostics:** The Root Mean Square Error (RMSE) of the fit between the estimated and observed spectral mixtures is listed for selected points in the scene. The first column lists the location and planned percentage mix of DiDy for these points. (Actually physically measured mixes were 100%, 78.8%, 50.5%, 24.2% and 0.0%.).

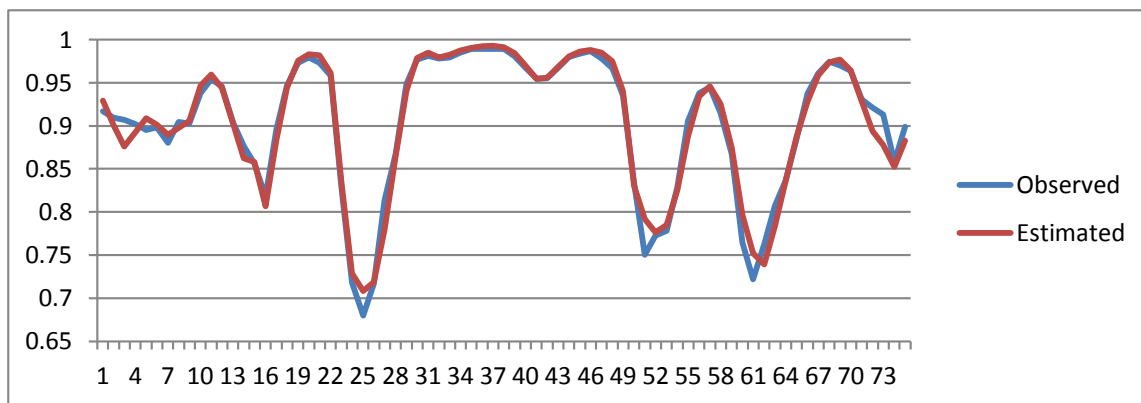
<u>(x, y)</u>	<u>DiDy %</u>	<u>FCLS</u>	<u>GKLS G=5</u>	<u>GKLS G=6</u>	<u>SSA</u>
(585,248)	100%	<b>0.0141</b>	0.0293	0.0369	0.0315
(400,228)	75%	0.0245	<b>0.0244</b>	0.0248	0.0351
(418,239)	75%	0.0183	<b>0.0177</b>	0.0324	0.0262
(405,251)	75%	<b>0.0179</b>	0.0196	0.0205	0.0255
(324 ,049)	50%	0.0357	0.0225	0.0222	<b>0.0129</b>
(313, 062)	50%	0.0359	0.0318	0.0390	<b>0.0118</b>
(328, 067)	50%	0.0333	0.0294	0.0533	<b>0.0223</b>
(223,314)	25%	0.0526	0.0380	0.0379	<b>0.0140</b>
(224,327)	25%	0.0495	0.0383	0.0384	<b>0.0186</b>
(246,330)	25%	0.0420	0.0289	0.0636	<b>0.0121</b>



(a) FCLS Method



(b) GKLS Method with Gamma  $G = 5$ .



(c) SSA Method

**Figure 9:** The observed and estimate estimated mixture (averaged) spectra of the 50/50% region using the FCLS, GKLS, and SSA methods.

Figure 9 shows the observed and estimated mixture (averaged) spectra of the 50/50% region using the FCLS, GKLS, and SSA methods. Visually, we can see both the GKLS (G=5) and SSA methods provide a better fit as compared to the FCLS method.

Table 2 lists the Root Mean Square Error (RMSE) of the fit between the estimated and observed spectral mixtures for selected points in the scene. Except for DiDy at 25% (0.242), the RMSE errors for FCLS were not considerably larger than the errors for the other methods. Yet we know FCLS is poorly predicting the known abundances for these samples. We conclude RMSE is not necessarily a good indicator of a methods accuracy to predict abundance. As far as the GKLS method, in most cases a gamma of G=5.0 provides a better fit than G=6.0. Recalling that G=5 provides a better prediction of abundance as compared to GKLS at the other values of gamma and also provides a smaller RMSE as compared to GKLS at G=6.0, we henceforth consider G= 5.0 to provide the best GKLS result.

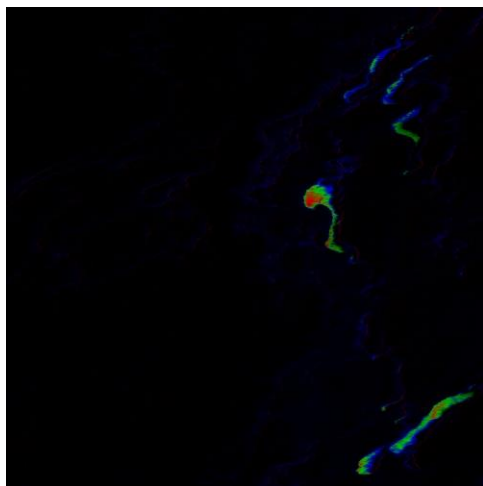
In Tables 1 and 2, the results also show (unfortunately) the automated implementation of the GKLS method, which attempts to select the most appropriate gamma based on achieving a minimum of the model's RMSE, was not as successful as the fixed gamma GKLS (G = 3, 4, 5, or 6) for estimating the correct abundance. This automated GKLS method attempts to select the most appropriate gamma based on achieving a minimum of the model's RMSE. We conclude the RMSE metric seems to respond to a mixture being linear or non-linear, but unfortunately it is not a reliable metric to determine the degree of non-linear behavior. RMSE could not be used to achieve the most accurate estimate of abundance and was not effective for implementing an automated GKLS.

The results for Experiment 2 (the Airborne Experiment) are shown in Figures 10 to 12 and Table 3 and 4. Figure 10 shows an RGB composite abundance map and five single-layer abundance maps for the fixed-gamma GKLS method. The composite abundance map shows visually the method's response to three thicknesses of oil: the R layer depicts the abundance of 4.0 mm. thick oil; the G layer depicts the abundance of 1.85 mm. thick oil; and the B layer depicts the abundance of 1.00 mm. thick oil. The single-layer abundance maps show visually the results for five of the six endmembers. Due to space limitations and the small size of these photos in the report, it is difficult to see the detailed patterns, but it's possible to get the sense of the behavior the method. We expect oil to spread out in a certain pattern from thick to thin, which is indeed what happens. In the original photos of the three methods (not shown), there is little difference in the patterns observed between the GKLS and SSA methods. However, there is a notable difference in the patterns seen for the 4.0 mm thick oil between the FCLS and GKLS (or SSA) methods.

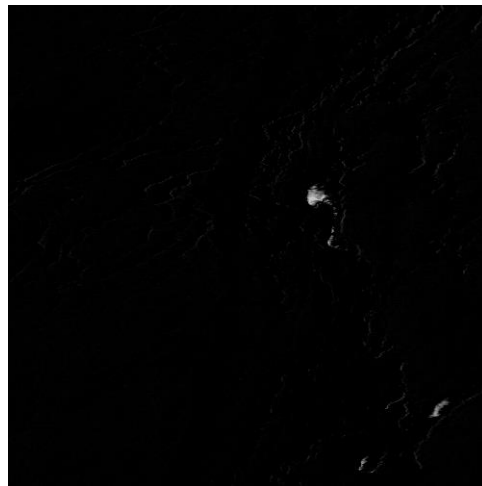
Figure 11 shows the single-layer abundance maps of the 4.0 mm. thick oil for the FCLS and fixed-gamma GKLS (G=5) method. The FCLS method shows much of the same large area of thick oil in the central-right side of the scene as the GKSL (and SAA, not shown). However, these maps also show a notable difference in pattern between the FCLS and GKLS methods, as just mentioned above. The linear method (FCLS) indicates a significant amount of 4.0 mm. thick oil throughout the scene in areas that would not be expected. Of course, we don't have ground truth, so it's always possible, but physically we expect the oil to spread out in a certain pattern from thick to thin. There is isolated 4.0 mm oil scattered throughout the area in the FCLS photo not seen in the GKLS photo. Neither the GKLS nor the SSA methods indicate much 4.0 mm. thick oil in these regions.

Tables 3 and 4 show the abundance results and model diagnostics, respectively, for the three methods at selected locations of the AVIRIS scene. In Table 3, the abundance estimates are not very different for the first three locations. Interestingly, the model diagnostics (RMSE) in Table 4 indicate the mixtures are likely to be linear at these locations. The abundance estimates listed in Table 3 for the last location (x,y) = (441, 441) are in agreement with Figure 11 and the smaller RMSE value for the GKLS at this location indicate mixture is likely non-linear.

Figure 12 shows plots of the observed and estimated spectra at the location (x,y) = (441, 441). The fits look good, visually, except for a significant difference in the F2 diagnostic region for the FCLS method and in the F1 diagnostic region for the SSA method. However, overall the fits are quite good considering the magnitudes of the data. The larger values of RMSE listed for the SSA method are likely a consequence of the magnitude of the SSA spectra as compared to the normalized reflectance values used in the FCLS and GKLS methods (see range of the y-axis in Figure 12).



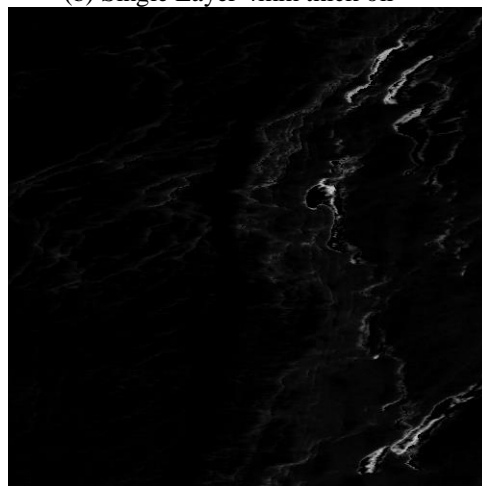
(a) RGB Layer: R=4mm, G=1.85m, R=1.0mm



(b) Single Layer 4mm thick oil



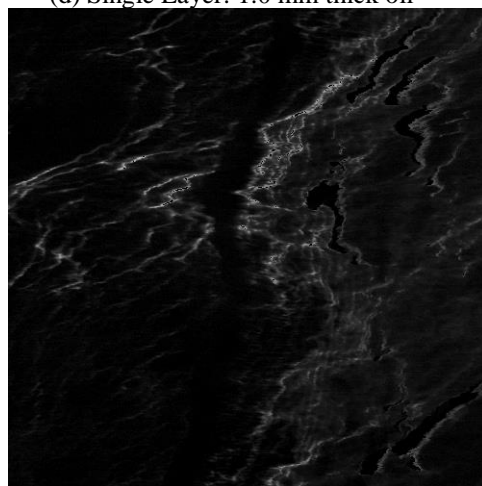
(c) Single Layer: 1.85 mm thick oil



(d) Single Layer: 1.0 mm thick oil



(e) Single Layer 0.5 mm thick oil



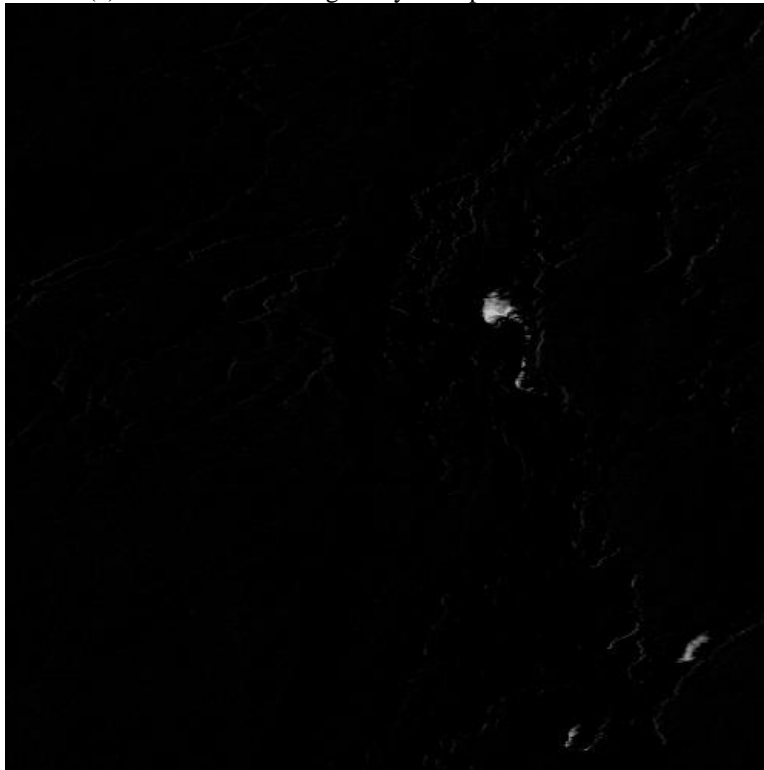
(f) Single Layer: Orange oil

**Figure 10:** RGB and Single-Layer Abundance Maps using the GKLS method are shown for 5 of the 6 endmembers (water endmember not shown).





(a) FCLS Method: Single Layer Map for 4 mm thick oil.



(b) GKLS Method (G=5): Single Layer Map for 4 mm thick oil.

**Figure 11:** Single-Layer Abundance Maps of the 4.0 mm thick oil are shown for the FCLS and GKLS (G=5) methods.

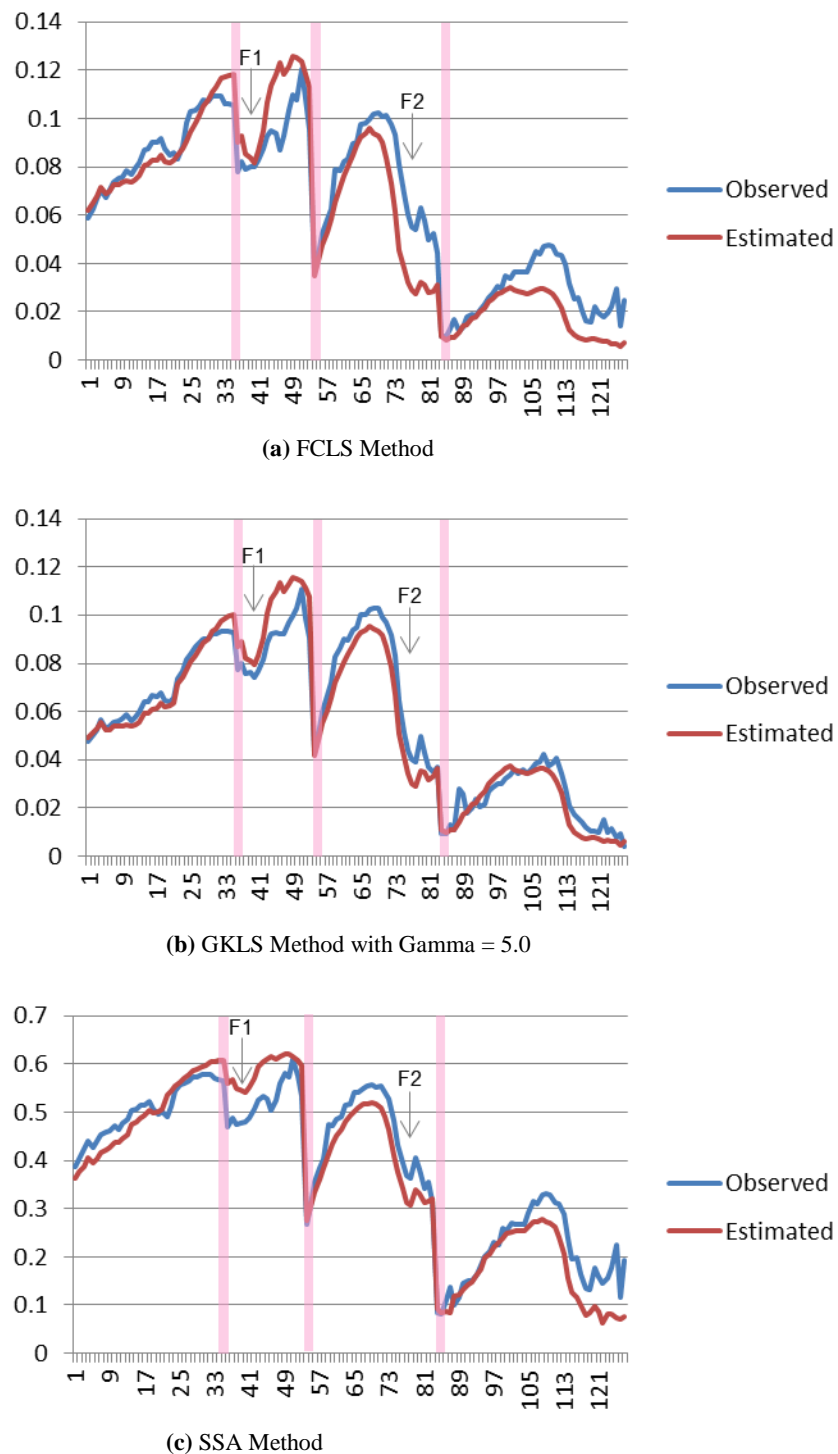


**Table 3. Abundance Results:** Estimated abundance values of the methods are listed for selected locations (points) of the AVIRIS oil scene.

		Lab	Lab	Lab	Lab	Img	Img
	Location	DWH Oil	DWH Oil	DWH Oil	DWH Oil	(546, 481)	(1154, 112)
<u>Method</u>	<u>(x,y)</u>	<u>4.0 mm</u>	<u>1.85 mm</u>	<u>1.0 mm</u>	<u>0.5 mm</u>	<u>Orange</u>	<u>Water</u>
FCLS	(317, 202)	0.965	0.000	0.000	0.000	0.000	0.035
GKLS=5	(317, 202)	0.956	0.000	0.000	0.000	0.000	0.044
SSA	(317, 202)	0.951	0.000	0.000	0.000	0.000	0.049
FCLS	(318, 202)	0.988	0.000	0.000	0.000	0.000	0.012
GKLS=5	(318, 202)	0.973	0.000	0.000	0.000	0.000	0.027
SSA	(318, 202)	0.965	0.000	0.000	0.000	0.000	0.035
FCLS	(323, 203)	0.966	0.043	0.000	0.000	-0.009	0.000
GKLS=5	(323, 203)	0.956	0.044	0.000	0.000	0.000	0.000
SSA	(323, 203)	0.952	0.048	0.000	0.000	0.000	0.000
FCLS	(410, 375)	0.133	0.000	0.171	0.000	0.187	0.508
GKLS=5	(410, 375)	0.137	0.000	0.230	0.052	0.201	0.380
SSA	(410, 375)	0.177	0.000	0.082	0.339	0.108	0.294
FCLS	(441, 441)	0.119	0.000	0.142	0.000	0.293	0.445
GKLS=5	(441, 441)	0.073	0.000	0.181	0.094	0.303	0.349
SSA	(441, 441)	0.075	0.000	0.000	0.460	0.177	0.288

**Table 4. Model Diagnostics:** The Root Mean Square Errors (RMSEs) computed for the fit between the observed and estimated mixed spectra are listed for selected locations of the AVIRIS oil scene.

Location	RMSE		
<u>(x,y)</u>	<u>FCLS</u>	<u>GKLS G=5</u>	<u>SSA</u>
(317, 202)	<b>0.0060</b>	0.0064	0.0274
(318, 202)	<b>0.0070</b>	0.0073	0.0286
(323, 203)	<b>0.0026</b>	0.0027	0.0133
(410, 375)	0.0092	<b>0.0073</b>	0.0287
(441, 441)	0.0132	<b>0.0118</b>	0.0484



**Figure 12.** Plots of the observed and estimated spectra at the selected location  $(x,y) = (441,441)$ . The result for the FCLS model is shown in (a), the result for the GKLS method at a fixed gamma of  $G=5.0$  is shown in (b) and the result for the SSA (ALBEDO) method is shown in (c). F1 and F2 are diagnostic oil features located at 1.2 microns and 1.73 microns, respectively. Bad bands have been removed in the plots. Artificial discontinuities are apparent for three of the four bad-band regions that were removed, at horizontal coordinates 35, 55, and 83. The discontinuities are indicated by the pink-colored strips.

## 5. CONCLUDING REMARKS

This study has investigated the use of a Generalized Kernel Least Squares (GKLS) method applied to reflectance data, and a Single Scattering Albedo (SSA) method for non-linear mixture analysis. In the case of the SSA method, a Fully Constrained Least Squares (FCLS) method is applied to data that has been converted from reflectance space to SSA space. Our baseline method was the FCLS method applied to reflectance data. Our hypothesis is that, for intimate (non-linear) mixtures, both of these methods will provide improved modeling and abundance estimates as compared to the baseline FCLS method. Two scenarios (laboratory and aerial) are tested on different scenes and hyperspectral sensors (Pika II and AVIRIS).

Overall the results for Experiment 1 indicate the FCLS method has a poor capability for modeling intimate mixtures. In contrast, both the GKLS and SSA methods do a much better job. Whether or not one is better than the other is not conclusive. However, we conclude that our hypothesis is confirmed and that both of these methods provide a better estimate of abundance for mixtures exhibiting nonlinearity. For the laboratory experiment of known abundance quantities, the SSA and GKLS methods responded well to the nonlinearity present in a mixture of materials and provided better estimates of abundance than the linear FCLS method for the DiDy and Lime.

The GKLS parameter “gamma” determines the degree of non-linear behavior exhibited by the GKLS method and affects its accuracy for estimating abundances. The automated GKLS method attempts to select the most appropriate gamma based on achieving a minimum of the model’s RMSE. We conclude the RMSE metric seems to respond to a mixture being linear and non-linear, but unfortunately it is not a reliable metric to determine the degree of non-linear behavior. It could not be used to achieve the most accurate estimate of abundance.

For mixtures known to be non-linear: A fixed gamma implementation of GKLS with  $G = 5$  or  $6$  provides a good estimate of abundance. The fixed gamma GKLS and the SSA methods can be computed in approximately the same amount of time and provide approximately the same accuracy for estimating abundances. The automated GKLS was much slower to compute and did not achieve better accuracy.

In Experiment 2, the aerial experiment over water (Deepwater Horizon Oil Spill), the SSA and GKLS methods performed equivalently and appeared to respond better to the non-linear mixtures of oil thicknesses in the water than the linear. Specifically, the linear method appears to be predicting too much of the thicker oils. This can only be determined qualitatively because of the lack of ground truth. Additional analysis should be performed on scenes containing a variety of known nonlinear mixtures that are well documented with ground truth.

## 6. ACKNOWLEDGEMENTS

Thanks are given to Dr. Roger Clark and Eric Livo (USGS) for the data provided in the second experiment, including the laboratory and field measurements, as well as the AVIRIS imagery.

**Approved for Public Release; Distribution Unlimited. 15-2513**  
**©2015 The MITRE Corporation. ALL RIGHTS RESERVED.**

## 7. REFERENCES

- [1] Adams J., Smith M., and Johnson P., "Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 Site," *J. Geophysical Research*, 91, (1986).
- [2] Boardman J., "Automating linear mixture analysis of imaging spectrometry data," *Proceedings of the International Symposium on Spectral Sensing Research (ISSSR)*, San Diego, CA. (1994).
- [3] Rand R.S., "A physically-constrained localized linear mixing model for TERCAT applications," *Proceedings of the SPIE Aerosense*, Orlando, FL., (2003).
- [4] Rand R.S., "Automated classification of built-up areas using neural networks and subpixel demixing methods on multispectral/hyperspectral data," *Proceedings of the 23<sup>rd</sup> Annual Conference of the Remote Sensing Society Remote Sensing Society (RSS97)*, Reading, United Kingdom, (1997).
- [5] Rand R.S., "Exploitation of hyperspectral data using discriminants and constrained linear subpixel demixing to perform automated material identification." *Proceedings of the International Symposium on Spectral Sensing Research (ISSSR)*, Melbourne, Australia, (1995).
- [6] Rand R.S. and Keenan D.M., "A Spectral Mixture Process Conditioned by Gibbs-based Partitioning," *IEEE Transactions on Geoscience and Remote Sensing: Special Issue on the Analysis of Hyperspectral Image Data*, 39(7), (2001).
- [7] Heinz D.C. and Chang C.-I., "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, 39(3), 529-545 (2001).
- [8] Montgomery D. and Peck E., [Introduction to Linear Regression Analysis], 2<sup>nd</sup> Ed., Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, (1992).
- [9] Hapke, B., (1993). *Theory of Reflectance and Emittance Spectroscopy*. Cambridge University Press, 455 p
- [10] Herzog S.G. and Mustard J.F., "Reflectance spectra of five component mineral mixtures: Implications for mixture modeling," *Lunar and Planetary Science XXVII*, 27, 535– 536 (1996).
- [11] Kwon H. and Nasrabadi N.M., "Kernel matched subspace detectors for hyperspectral target detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28 (2), 178-194 (2006).
- [12] Camps-Valls and Bruzzone L., "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, 43(6), pp.1351-1362, (2005).
- [13] Scholkopf B and Smola A.J., [Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond], The MIT Press, Cambridge, MA, (2002).
- [14] Broadwater J.B., Chellappa R., Banerjee A., Burlina P., "Kernel fully constrained least squares abundance estimates," *Proceedings of the IEEE International Geoscience and Remote Symposium (IGARSS 2007)*, Barcelona, Spain, 4091-4044 (2007).
- [15] Broadwater J.B. and Banerjee A., "A generalized kernel for areal and intimate mixtures," *Proceedings of the IEEE WHISPERS '10*, Reykjavik, Iceland, (2010).
- [16] Broadwater J.B. and Banerjee A., "Mapping intimate mixtures using an adaptive kernel-based technique," *Proceedings of the IEEE WHISPERS '11*, Lisbon, Portugal, (2011).

- [17] Broadwater J.B. and Banerjee A., "A comparison of kernel functions for intimate mixture models," *Proceedings of the IEEE WHISPERS '09*, Grenoble, France, (2009).
- [18] Mustard J.F. and Pieters C.M., "Photometric phase functions of common geologic minerals and application to quantitative analysis of mineral mixture reflectance spectra," *Journal of Geophysical Research*, 94, 13619-13634 (1989).
- [19] Brent R.P., [Algorithms for Minimization without Derivatives], Prentice-Hall, Englewood Cliffs, NJ, (1973).
- [20] Resmini, R.G., Graver, W.R., Kappus, M.E., and Anderson, M.E., (1996). Constrained energy minimization applied to apparent reflectance and single-scattering albedo spectra: a comparison. *Proceedings of the SPIE: Hyperspectral Remote Sensing and Applications*, Sylvia S. Shen, ed., Denver, Colo., August 5-6, v. 2821, pp. 3-13, doi:10.1117/12.257168.
- [21] Resmini, R.G., (1997). Enhanced detection of objects in shade using a single-scattering albedo transformation applied to airborne imaging spectrometer data. *The International Symposium on Spectral Sensing Research*, San Diego, California, CD-ROM, 7 p.
- [22] Nascimento, J.M.P., and Bioucas-Dias, J.M., (2010). Unmixing hyperspectral intimate mixtures. *Proceedings of the SPIE*, v. 7830, doi:10.1117/12.8651188, 8 p.
- [23] Rand R.S., Banerjee A., Broadwater J., "Automated endmember determination and adaptive spectral mixture analysis using kernel methods," *Proceedings of SPIE*, Optics and Photonics, San Diego, CA, August (2013).
- [24] Resmini R.G., Rand R.S., Allen D.W., Deloy C.J. "An analysis of the nonlinear spectral mixing of didymium and soda lime glass beads using hyperspectral imagery (HSI) microscopy" *Proceedings of SPIE 9088*, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XX, 9088OZ (June 13, 2014), doi:10.1117/12.2051434 (<http://dx.doi.org/10.1117/12.2051434>), 15p.
- [25] Gao B.-CI., Montes M.J., Ahmad Z., and Davis C.O., "Atmospheric Correction Algorithm for Hyperspectral Remote Sensing of Ocean Color from Space" *Applied Optics*, 39(6), (2000).
- [26] Gao B.-CI. and Davis C.O., "Development of a line-by-line-based atmosphere removal algorithm for airborne and spaceborne imaging spectrometers" *Proceedings of SPIE*, 3118, 132-141 (1997).
- [27] Clark R.N., Swayze G.A., Leifer I., Livo K.E., Kokaly R., Hoefen T., Lundeen S., Eastwood M., Green R.O., Pearson N., Sarture C., McCubbine I., Roberts D., Bradley E., Steele D., Ryan T., Dominguez R., "A Method for quantitative mapping of thick oil spills using imaging spectroscopy," *U.S. Geological Survey Open File Report*, (2010).
- [28] Rand R.S., Clark R.N., and Livo K.E., "Feature-based and statistical methods for analyzing the Deepwater Horizon oil spill with AVIRIS imagery," *Proceedings of SPIE*, Optics and Photonics, San Diego, CA, August (2012).
- [29] [http://www.resonon.com/imagers\\_pika\\_iii.html](http://www.resonon.com/imagers_pika_iii.html) (last accessed on Dec. 3, 2013).
- [30] We have also used an Edmund Optics Gold Series 1.0X telecentric lens that gives ~8  $\mu\text{m}/\text{pixel}$ . However, data at such a high spatial resolution were not required for the analyses reported upon here.
- [31] Note: References are made to certain commercially available products in this paper to adequately specify the experimental procedures involved. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that these products are the best for the purpose specified.
- [32] Schott, J.R., (2007). *Remote Sensing: The Image Chain Approach*, 2<sup>nd</sup> ed. Oxford University Press, New York, 688 p.
- [33] Photographs in Figures 1 and 2 taken by, owned by, and courtesy of co-author Dr. David W. Allen of NIST.